Getting the Data In: Three Year Experience With A Pediatric Electronic Medical Record System

Isaac S. Kohane, Division of Endocrinology, Children's Hospital and Harvard Medical School, Boston, MA

The Clinician's Workstation (CWS) has provided the full-functionality of an on-line electronic patient record for outpatient pediatric clinics over the past 3 years. The implementation of the CWS built upon a substantial effort in integration of data from various sources. This paper addresses the subsequent design issues which had to be resolved in order to enable both physician and transcriptionist-driven data entry and retrieval, notably selecting a feasible mixture of controlled vocabulary and free text. Some of the consequences of these design decisions on clinical care, clinical education, clinical and basic research are reviewed with examples from the last three years.

INTRODUCTION

Over the past 5 years, the trend towards implementing a fully-electronic medical record has accelerated considerably. Comprehensive electronic medical record systems (EMRS) have been discussed for decades [1], but comprehensively implemented in relatively few sites (e.g. Regenstrief [2]. HELP,[3], TMR [4]). There have also been several efforts to develop workstation-based, graphics-intensive on-line patient charts (e.g. PWS at Hewlett Packard/Stanford, [5]). These efforts only addressed the specific requirements of pediatric EMRS tangentially. I describe here the design and performance of a pediatric EMRS—the Clinician's Workstation (CWS)—built upon a "client-server" architecture. The CWS has been in operation for three years, first in the Division of Endocrinology and more recently in the Divisions of Nephrology and Nuclear Medicine.

The data-integration efforts that led to the development of the CWS will only be briefly touched upon as these have been published elsewhere[6,7]. Rather this paper focuses on the task that McDonald et al. [8] have termed "the difficult side of medical record systems," namely data acquisition and in particular acquisition of data from clinicians. This focus includes the tradeoffs that have made between clinician acceptance, and the requirements for controlled, coded vocabularies. It also describes the technological solutions and organizational solutions required to implement these tradeoffs. The consequences of our* design choices are illustrated by providing a few illustrative examples of how the CWS can be used to 1) improve clinical efficiency, 2) enable clinical and basic science research and 3) quantify some aspects of clinician performance. Those aspects of the CWS that are specific to the practice of pediatrics will be emphasized.

*Several members of the Division of Endocrinology, particularly Drs. Majzoub and Crigler, were influential in the design process. This work was supported, in part, by the Charles Hood Foundation.

Summary of Current CWS Functionality.

The current version of the CWS is implemented on Macintosh computers networked to the Integrated Hospital Information System (IHIS). The IHIS has as its centralized data repository, an Oracle database stored on several Digital Equipment Corporation VAX computers (the "VAX Cluster"). The CWS retrieves and displays all pertinent administrative, financial and clinical data residing on the VAX Cluster. These data include: demographics, visit history with associated procedure and diagnostic codes, inpatient pharmacy orders, inpatient laboratory studies which are entered into the IHIS through other departmental applications (e.g. the Cerner laboratory system). Users of the CWS enter additional clinical documentation into the IHIS through the CWS interface. These data include: problem lists, patient-provider relationships, bedside measurements, history, past medical history, family history, review of systems and other components of clinic notes or letters to referring physicians. Access to this information is controlled by assigning data access/modification privileges to various provider roles. The CWS serves to maintain *all* clinical data/documentation of patients seen by *all* clinicians in each participating clinic. Data displays are designed to follow the metaphor of the paper chart as possible but employ other metaphors where appropriate.

The client program on the Macintosh computers was written in Hypercard. Transactions with the Oracle data-base are communicated through SQLNet protocols (Oracle Corporation) running over a hospital-wide ethernet network.

DESIGN CONSIDERATIONS

Goals

At the outset we set the bounds on the technological solutions we envisaged by insisting that we meet, within a 1.5 year implementation project plan, the following goals.

1. The CWS would be universally used within each clinic in which it was deployed.
2. Clinicians should be able to rapidly enter free text and coded data items as part of the routine workflow in the care of their patients.
3. Clinicians who did not wish to perform data entry should be able to dictate the clinic visit into a tape recorder and have transcriptionists enter data. This alternate route should not provide less coded and objective information than the direct entry method.
4. Provide entry and display screens that are useful and familiar to pediatricians.
5. Provide sufficient accurately coded and quantified data to support automated clinical event monitors, clinical research, and outcomes research.
6. Avoid redundant data entry: From a single entry of data from a clinic visit, all other documentation (e.g. clinic notes, letters to referring physicians) should be generated.

7. Training requirements would have to be minimal in order to accommodate the large number of transient clinicians rotating through our clinics.

Constraints

Further we were constrained by the following limitations:

1. Voice-recognition and handwriting recognition technology was not mature to reliably and accurately encode terms for the relatively unconstrained domain of pediatric histories.

2. We could not guarantee that each clinician could always have immediate access to a workstation for data entry. Each clinic was provided with between 5 to 7 Macintosh computers networked to the IHIS. This limitation has since been made moot by the collapse of hardware costs but was a significant consideration in our original design.

3. In the late 1980's, off-the-shelf client authoring tools for SQL-compliant data-bases were scarce and had limited capabilities.

4. Post-hoc parsing of free text (e.g. [9]) is not sufficiently accurate to achieve goal 5 (above).

IMPLEMENTATION

The approach we took for the task of clinician-driven data entry was three-fold: 1) an electronic form was created in the CWS for data-entry 2) clinicians were given a variety of paper-based equivalents to the electronic form and 3) an ongoing program of clinician feedback and software modification was implemented. These three components are described below.

Electronic Form

The purpose of the electronic form is to enable data entry to be performed at very close to the speed of unrestricted typing. The user of the form tabs from field to field within the form and is only prompted upon detection of potential data entry errors.

If a patient has already been seen once, then the CWS automatically retrieves the following items which therefore do not have to be entered by the clinician: patient/parent's address, the address of referring clinician(s) and laboratory studies at the time of the visit. Each clinic using the CWS can define those data elements that they wish to be encoded for later systematic analysis. In the Endocrine clinic these include standard anthropometry (e.g. height, weight, arm span) and sexual development information (e.g. testicular size, Tanner staging). Within the form, a field is created for each such data element in the order that the clinic providers are accustomed to. As the provider enters the values of these data, a clinical data extraction program associated with each field is triggered. The default program, which can be customized, stores the content of the field in a clinical data table on the server. Bookkeeping details such as the date the clinical finding was observed, the time of data entry and the code of the data type are automatically determined by the data extraction program. For some data elements, the default data extraction program has to be modified for specialized data validation. For example when one of the testicular short axis measurements are entered, the data extraction program checks whether this measurement is less than the long axis. If not, it offers to switch the two measurements. We have summarized below the three most important classes of coded data elements: physical exam bedside measurements, patient:provider relationships and problem lists. The non-coded fields in the electronic form merely tag the sections of the unrestricted or "free" text of the clinical note (e.g. family history) so that the program which generates the letters to referring physicians or the clinical note for the chart can manipulate and position the text fragments appropriately.

Physical Exam. The tempo and pattern of growth and development of children is among the most sensitive measures of health. Many disorders can be first detected through careful inspection of the growth and development data routinely acquired during the course of regular pediatric visits [10]. We have consequently encoded several standards for the progression of growth and development parameters (e.g. blood pressure, height, sexual staging) with age. Where possible we have encoded longitudinal standards obtained for the many distinct populations that pediatricians will follow (e.g. Turner's Syndrome, Down's Syndrome, late and early puberty).

These encoded standards serve to improve data validation (e.g. through automated identification of implausible changes in standard deviation score for height) by the default data extraction program of the electronic form. They also have enabled us to generate data displays that are familiar to pediatricians such as the growth chart in Figure 1. This chart was generated using one [11] of several standards for growth and another standard for predicting adult height from bone age [12].

Patient:Provider Relationships. Particularly in tertiary-care centers, many providers participate in the health-care of each patient. Furthermore, a large fraction of these same providers (fellows and residents) will only follow these patients for a few years. To prevent unintended gaps in patients care, we chose to explicitly enter patient:provider relationships as part of the data entry process. Every clinic visit document has a primary signatory and a large subset of them also have a secondary signatory (if the attending physician sees the patient with a fellow or resident). Clinicians "sign" the documents generated from the electronic form by entering their unique provider identification number. This populates a patient:provider relationship table on the server data-base in the IHIS. The table also stores the time that the relationship was established and the role that the provider serves for the patient (e.g. supervisory, primary or research).

Problem Lists. Problem lists serve to quickly summarize a patient's course. They can also serve to identify subpopulations of clinic patients of relevance to clinical or basic science research or outcomes studies. As part of the data entry process, each patient is assigned one or more problems (e.g. autoimmune thyroiditis) from a vocabulary that is specific to each clinic. Each vocabulary term is classified in a nosology to permit aggregation of these subpopulations (e.g. to find all patients with thyroid disease). Unfortunately most standard controlled vocabularies do not provide sufficiently fine-grained descriptors for all the pathophysiological disorders we would like to capture. Therefore, for each clinic using

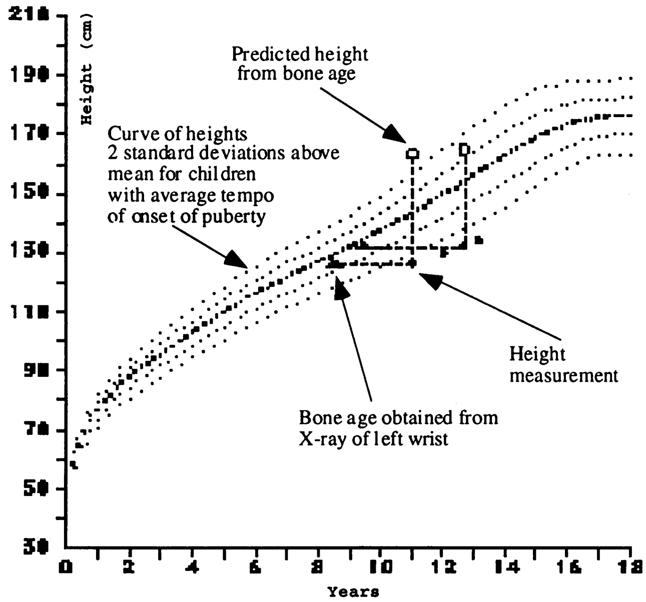


Figure 1: Growth Chart

the CWS, the clinicians must arrive at a consensus vocabulary for problems. These vocabularies are periodically updated. We are now considering requiring for each such consensus vocabulary, a set of mappings to the ULMS vocabulary[13], which even though not as fine-grained is standardized.

Paper Forms

As noted in the design considerations section, we did not want participation in the CWS data entry to be limited by access to workstations or clinician resistance to use of computers. Therefore, we created paper equivalents of the electronic form that clinicians use for taking notes during the course of the patient visit. For each field in the electronic form there is an equivalent labeled area on the paper form in identical order.

Many clinics at Children's Hospital still use the traditional paper chart, therefore the paper form is backed with "carbonless" pressure-sensitive paper so that a duplicate copy of the notes can be left in the paper chart before the electronic version is generated. If the clinician does not directly enter the documentation of the clinic visit into the CWS, a transcriptionist will enter the documentation either using the paper form or a taped, dictated summary. Any clinician dictating the documentation for a clinic visit follows the same order that the data appears on the electronic and paper forms. Index cards listing the data elements in order have been provided for this purpose.

Clinician Participation

Clinician acceptance of the CWS was recognized as the principal hurdle from the outset. Consequently, physician and nursing staff in the target clinics were appraised of major design decisions at regular intervals during the design process. The appearance and function of the electronic and paper forms have undergone several revisions since the onset of the design process.

We found that the transcriptionists and administrative staff were among the more frequent users of the CWS and that the success of the entire project depended critically on their ability to use the client program efficiently. Bottlenecks became rapidly apparent both in the automated auditing of the transcription process performed by the CWS and in the comments coming from the administrative staff. These comments led to repeated streamlining and simplification of the data entry process as well as automation of mundane but onerous ancillary tasks (e.g. addressing envelopes, creating address lists of patients followed by a particular physician).

Our approach to user participation in the design process has dictated an incremental, clinic-by-clinic adoption of the CWS rather than attempting a hospital-wide implementation. Given the lessons learned during the course of its deployment and the varying requirements of each clinic, this seems to have been a prudent course.

RESULTS

Since its first deployment in July of 1991, the CWS data-base has accumulated the records of 3100 patients (i.e. 100% of patients seen in the clinic). Excluding reports generated by other departmental applications (e.g. radiology, pathology which are accessible through the same CWS interface) 6500 visit forms were completed. In the process, 38,000 individually coded clinic measurements were automatically entered into the data-base as well 3400 problems (using the clinics' controlled problem list vocabulary). As the first 2.5 years of its deployment were restricted to a single clinic, we anticipate rapid growth in these numbers in the near future.

In this section we describe the impact of the CWS deployment with a few selected examples that we have organized into four rubrics that we have found to represent important uses of the CWS.

Clinical Care

The most obvious consequence of implementing the CWS is the availability of the patient record. Whereas previously, at best records were missing or misplaced for approximately 5% of patient visits, we now have immediate access to documentation on all visits to clinics where the CWS has been deployed. With the hospital-wide ethernet network, these records can be viewed, with proper authorization, throughout the institution.

In the Division of Endocrinology, there are 19 physicians who use the CWS and approximately 20 visiting physicians (fellows from other institutions and housestaff from Children's Hospital) per year. Electronic data entry is performed by only 15% of clinicians whereas 20% submit the paper forms with handwritten entries and 65% submit taped dictations.

By selecting 100 clinic visits from immediately prior to the deployment of the CWS and 100 one year after its deployment, the time from a patient visit to the sending of a letter to the referring physician declined from approximately 3 weeks to 2.1 weeks (the null hypothesis of the mean follow-up time in the two periods being equal was rejected with $p < 0.05$). Many factors may have contributed to this trend other than the CWS client-server application. These include the effect that the paper forms may have had in standardizing the data acquisition behavior of clinicians. It could also be explained in part by other factors such as changes in administrative staff.

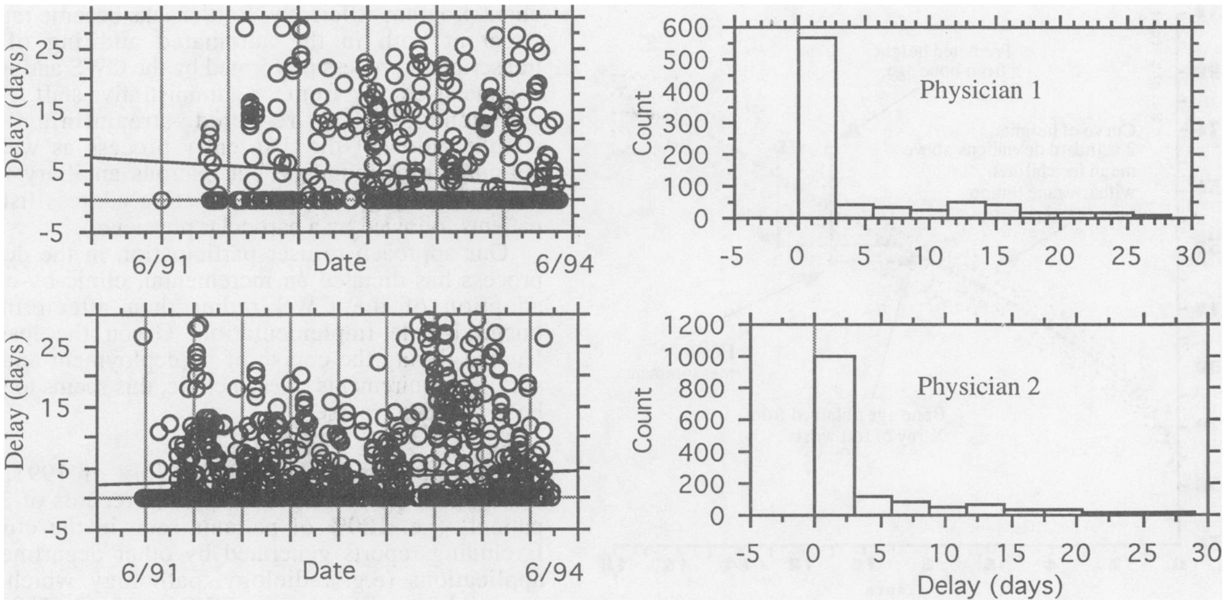


Figure 2: Delay Between Patient Visit and Completion of Documentation for Two Physicians.

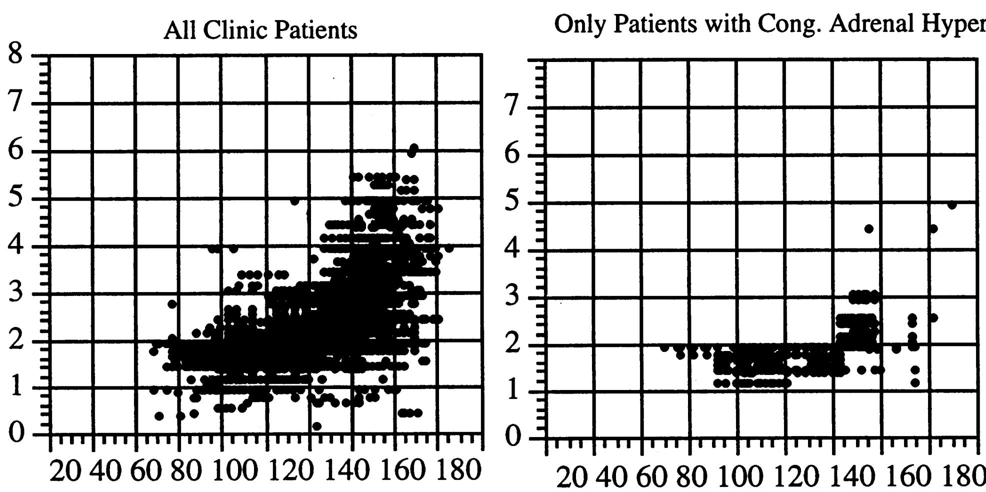
Nonetheless, in the absence of controlled trials, these results are encouraging. We note however, that the model of the CWS use is designed for tertiary care clinics. In other settings, such as high-volume primary-care clinics, it may present a suboptimal model for entry of clinician-derived data.

Quality Assurance

During the process of generating a clinical document such as a clinic note or a letter to a referring physician, the CWS enters a large amount of bookkeeping detail which has enabled us to implement several quality assurance programs. This includes the identity of the clinicians involved, the transcriptionists, the date of the clinic visit, the date the document was first created, the date it was last modified and the date "published" (at which point it can no longer be edited). Although we are still in the process of picking those monitors or filters that will be the most useful, we illustrate here (Figure 2) one potentially interesting application. In these graphs we illustrate the delay between the date the patient was seen and the date the letter to the referring physician was completed. One of the two physicians clearly has a lighter clinical load (spends a greater percentage of time in basic research) and is less prompt in completing the documentation although the plotted regression line shows some steady improvement over the past three years. These plots do not control for patient case mix.

Clinical Research

The CWS has already enabled several clinical research projects that would otherwise have been prohibitively labor intensive, to get underway. This includes a study of the dose-response relationship for growth hormone in growth hormone-deficient patients [14], and a review of outcome predictors in patients



Plots of Long Axis of Testicle (cm) vs. Height (cm)

Figure 3: Graphs for Clinical Research from the CWS

with non-classical adrenal hyperplasia (in progress). As clinical data continues to accumulate as a side-effect of routine health-care documentation, we anticipate that not only will we be able to quickly generate cost-effective clinical studies but we will be able to generate new, comprehensively documented, standards for a variety of pediatric parameters (e.g. problem-specific growth curves). Again, only to illustrate the capabilities of the CWS, we show (in Figure 3) a graph of testicular dimension (long axis) plotted against total standing height. These parameters are plotted for both for the entire clinic population and for also for those patients with congenital adrenal hyperplasia. Although somewhat whimsical in their specifics, these graphs demonstrate how the various kinds of coded data (in this case bedside measurements and problem lists) stored by the CWS can be used to aid clinical research.

The emergence of reasonably robust interapplication protocols on personal computers (as the CWS is implemented on the Apple Macintosh we have used the AppleEvents protocol) has enabled us to display, in real-time, within commercial graphical applications (e.g. Deltagraph from Deltapoint or Excel from Microsoft) the results of queries initiated in the CWS client application. Figure 3 was generated in this way. The interapplication protocols have also enabled us to automatically route clinical alerts from the CWS to the electronic mail system.

Basic Research

We have found that the CWS can serve to generate a low-cost bridge between basic research and clinical practice. For example, for a collaborator interested in the specific gene defects leading to obesity, we were able to generate a list of all patients with a very high weight for height (using pediatric standards encoded in the CWS) and who did not have known CNS malignancies (obtained from the coded problem list). The CWS enabled another researcher to find a group of patients with combinations of neuroendocrine insufficiencies from which she identified a novel mutation of the Pit-1 gene [15].

CONCLUSION

The Clinician's Workstation's implementation has have largely met the goals that we set for ourselves in the design stage. Initial results over the first three years of deployment suggest that the specific combination of free text and controlled vocabularies we have chosen is effective in meeting these goals. Further, as the users of the CWS become more familiar in its use and as its data-base has become more substantial, we have begun to see it used for clinical productivity and research in ways we had not anticipated.

Nonetheless, five years after we implemented the first prototypes, some limitations have become apparent. The tools we chose to implement the client application have relatively poor performance and versatility as compared to the client-building tools available today. Also, the cost of high-performance hardware and the capabilities of system software have reached levels that make technologies such as pen or voice recognition potentially viable. If these technologies result in a much higher clinician acceptance of direct data entry, then a much larger portion of the visit documentation could be encoded in controlled vocabularies. The current CWS information infrastructure will permit us to explore the added value of these tools in the near future.

REFERENCES

1. Greenes, R. A.; Pappalardo, A. N.; Marble, C. W.; Barnett, G. O. Design and implementation of a clinical data management system. *Comput. Biomed. Res.* 1969, *2*, 469-485.

2. McDonald, C. J.; Blevins, L.; Tierney, W. M.; Martin, D. K. The Regenstrief Medical Records. *MD Computing* 1988, *5*, 34-47.

3. Pyror, T. A.; Gardner, R. M.; Clayton, P. D.; Warner, H. R. The HELP system. *Journal of Medical Systems* 1983, *7*, 87.

4. Stead, W. W.; Hammond, W. E. Computer-based medical records: The Centerpiece of TMS. *MD Computing* 1988, *5*, 48-62.

5. Tang, P. C.; Annevelink, J.; Fafchamps, D.; Stanton, W.; Young, C. Y. Physician's workstations: Integrated information management for clinicians. In: *Proceedings Symposium on Computer Applications in Medical Care.* P. D. Clayton, Eds., Washington, DC: McGraw-Hill, 1992:569-573.

6. Margulies, D.; McCallie, D. P.; Elkowitz, A.; Ribitsky, R. An integrated hospital information system at Children's Hospital. In: *Proceedings SCAMC.* Washington, DC: 1990:699-703.

7. Kohane, I. S.; David P. McCallie, J. A dynamically reconfigurable clinician's workstation with transparent access to remote and local databases. In: *First Annual American Medical Informatics Conference.* Snowbird, Utah: 1990:

8. McDonald, C. J.; Tierney, W. M.; Overhage, J. M.; Martin, D. K.; Wilson, G. A. The Regenstrief Medical Record System: 20 years of experience in hospitals, clinics and neighborhood health centers. *MD Computing* 1992, *9*, 206-217.

9. Salton, G. Development in automatic text retrieval. *Science* 1991, *253*, 974-980.

10. In *Nelson Textbook of Pediatrics*; 12 ed.; R. E. Behrman and V. C. Vaughan, Ed.; 1983; pp. 186-194.

11. Hamill, P. V. V.; Drizd, T. A.; Johnson, C. L.; Reed, R. R.; Roche, A. F. In *Vital Statistics Report*Rockville, MD, 1976; Vol. 25.

12. Greulich, W.; Pyle, S. *Radiographic Atlas of Skeletal Development of Hand and Wrist*; Stanford Press: 1959.

13. Sheretz, D.; Tuttle, M.; Blois, M.; Erlbaum, M. Intervocabulary mapping within the UMLS: the role of lexical matching. In: *Proceedings SCAMC.* R. Greenes, Eds., Washington, DC: IEEE Computer Society Press, 1988:201-206.

14. Kohane, I. S.; K.Faizan; Adjanee, N.; Najjar, S. S. Can cost effectiveness of growth hormone be improved? *Pediatric Research (Suppl.)* 1993, *33*, S51.

15. Cohen, L. E.; Wondisford, F. E.; Radovick, S. A novel mutation in the phosphorylation consensus sequence of the Pit-1 gene in a patient with dysregulation of prolactin and thyrotropin secretion. In: *76th Annual Meeting of the Endocrine Society.* Anaheim, CA: 1994:Abstract #6.